LogDoor User's Guide

LogDoor Files

LogDoor Task files

You specify the details of how LogDoor should process a log through LogDoor task files. Each LogDoor task file specifies an input log file to process, and output file and scheduling information. A LogDoor task file also holds a pointer to the last byte in the input log file which has been processed by the task. Finally the task file maintains a complete summary, on a day-by-day, site-by-site basis, of all the information it has processed to date. This information is displayed in the task status window and written, on a periodic basis, to the site-by-site summary reports.

- NOTE: Detailed day-by-day information is maintained only for the most recent 90 days. Information prior to that period (or from log entries that don't include a valid date) is grouped together in a category called "Old".

In many cases you will only need one LogDoor task file. If however, you have multiple log files (for instance one for your Web server and one for HomeDoor), or want to process the log file in different ways, you may want to set up a number of LogDoor task files. LogDoor can only have one task file open at a time. You can create a new LogDoor task file, or choose one that's been created previously, only if no other task file is currently opened. Use "Close Task" from the File menu to close the currently open LogDoor task (and terminate processing if any is ongoing). Use "New Task" to create a new task file, or "Open Task" to open one which has previously been created.

- IMPORTANT: Do not move a LogDoor task file while the task defined by that file is running.

As indicated above, the LogDoor task file contains all of the state information associated with the task's processing. This information is checkpointed out to disk on a periodic basis. If you wish to be sure this information is saved at a specific time, you can choose "Checkpoint Now" from the Task menu

- NOTE: LogDoor displays details of the task's processing state in the task status window. You can hide this window without affecting the task's processing by choosing "Hide Task Status Window" from the Windows menu, or simply by closing the window.

To remove an old site from the state information associated with the task, and from the task status window, select the site in the task status window and choose "Remove Site" from the Task menu. All state information about the site will be deleted, and no future log files or reports for the site will be written (unless a valid access to the site is encountered during subsequent processing).

- NOTE: Previous accesses to the site will continue to be included in overall totals for your server.
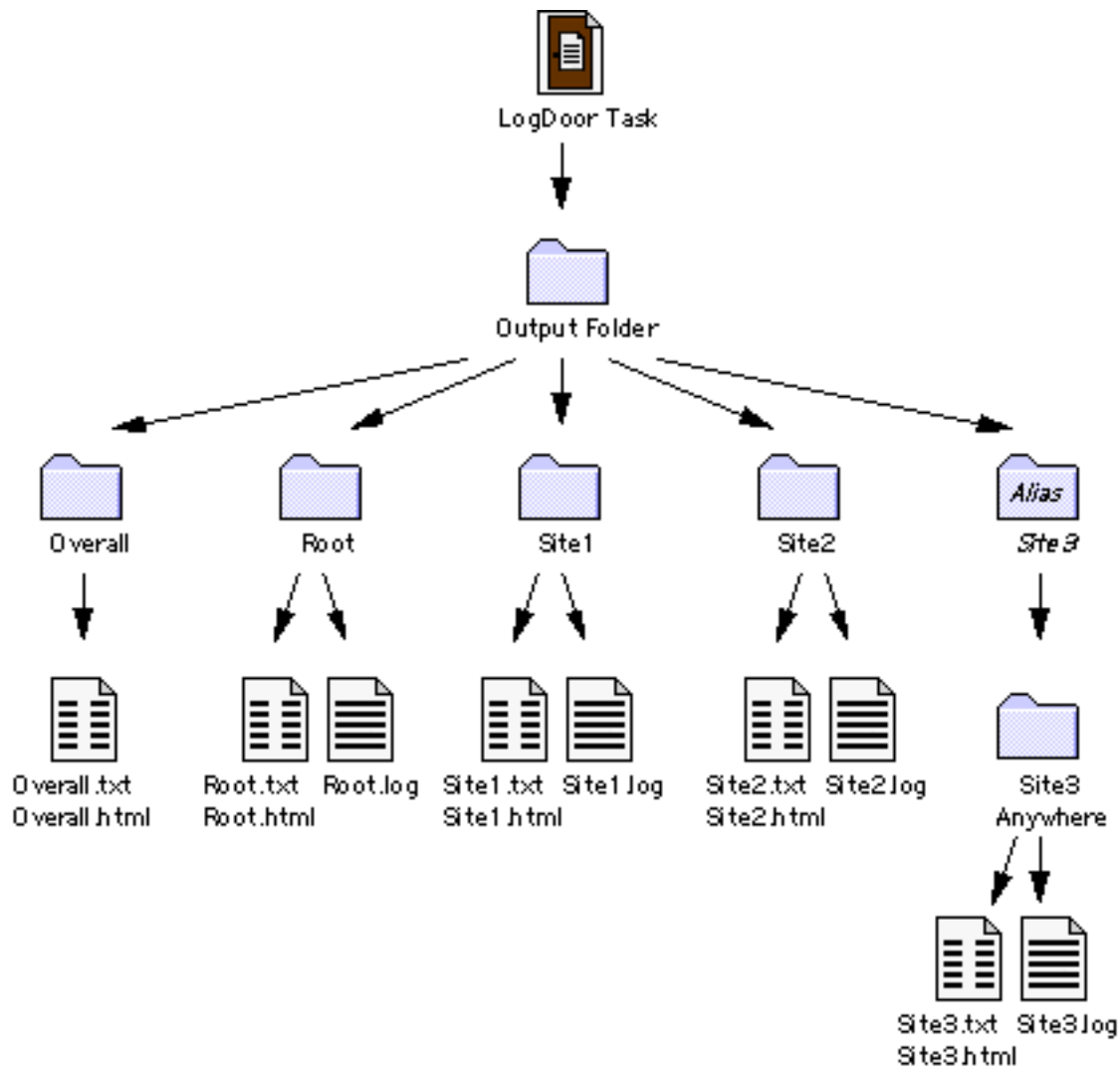
LogDoor's Output Files

Much of LogDoor's usefulness comes from the output files which it produces. This section provides details of those files.

LogDoor's output files are based on the "site" indicated in each entry of the input log file. LogDoor defines the site as the highest-level folder indicated in the URL field of the entry. If there is no folder in the URL, the site is defined as "root." Here are some examples:

```
URL                             Site

:Site1:default.html             Site1
:Site1:folder:file.html         Site1
:Site2:f1:f2:f3:file.html       Site2
:default.html                   Root
:file.html                      Root
```

- NOTE: The definition of a "site" in the HomeDoor log file is the same as in a Web server log file. HomeDoor logs simply add a host name to their URLs, which is ignored for purposes of site definition. See the section "HomeDoor-specific processing" for details.

The following figure presents an overall summary of the type and structure of a LogDoor task's output files.

As shown above, all of a task's output is written into the folder indicated to the task through the Output Files button or menu item. If Output Files is never selected, the task's output is written into a folder named "Output" in the same folder as the task itself.

- IMPORTANT: Do not move a task's output folder while the task is opened. If you do so, a new output folder at the previous location will be created the first time a new site is processed, and all new sites from then on will be put into that folder.

For each site that the task encounters in the input log, the task looks in its output folder for a folder with that site name. If one is not found, the task creates a new folder for the site. Either way, the task then writes its output files into that folder, creating new files as necessary. If the folder is actually an alias, the task writes its output files into the folder pointed to by that alias, allowing the files to be written, for instance, into folders at different places within your Web hierarhcy (see Appendix 1 for some examples).

- NOTE: A LogDoor task processes input log lines with error (ERR!) indications differently under certain conditions. Specifically, if an error line is encountered for a site not currently in the task's site list, the line is written to the root log file rather than creating a new output folder and log file, and the site is not added to the site list. This processing prevents the creation of a large number of entries and output folders for invalid sites (for instance sites with names that are mistyped by browsers).

A LogDoor task can create up to three files for each site that the task encounters in the input log:

- Site log file - produced as the input log file is being read by the task, this file includes every line of the input log file which indicates access to a URL for the site.

- Site processing summary (HTML format) - written out by the task on a periodic basis (or on demand), this file summarizes the processing performed by the task for the specific site.

- Site processing summary (text format) - the same information as the HTML file, but in tab-delimited text format.

You can use the Output Files dialog to indicate which of these files should be produced by a LogDoor task.

When writing output logs, the LogDoor task echoes any input log format line to each output log file, so that log post-processing applications will be able to correctly process LogDoor output logs. Site logs are always appended to and are never reset by a LogDoor task.

- NOTE: A LogDoor task opens one output log file for each site it encounters in the input log file, and keeps that file open until task processing is terminated. LogDoor will not, however, ever open more than 200 log files at any one time, so as to ensure that the Macintosh does not approach the MacOS's open file limit.

Site logs, like all three of LogDoor's output file types, are named based on the site name. LogDoor appends a user-specified suffix to the site name to form the output file name. You can specify the suffix in the Output Files dialog. The default suffix used for log files is ".log", so the name of a log file will generally have the format <site-name>.log.

In addition to indicating that a file is of a specific type, suffixes are useful for changing the names of LogDoor's output files at a particular time. For instance, you could use a suffix which included the month name, and change that suffix at the end of every month, causing LogDoor to write out new logs each month. If a suffix causes the resulting file name to have more than 32 characters, the site name will be truncated as necessary.

- NOTE: If you wish to export any of LogDoor's output files over the Web, do not include special characters (such as most punctuation other than period or hyphen) in the suffix. LogDoor will warn you if you have included a character which can not be used in file names over the Web.

The HTML and text-based summary reports are written by the LogDoor task on a periodic basis, as indicated in the Scheduling dialog box (see the "Scheduling" section). Summary reports can also be written at any time through the "Write Summaries Now" item in the Task menu. These summaries contain day-by-day totals of the task's processing in terms of the number of hits, bytes transfered and errors for pages within each site. Unlike the output log files, summaries are reset by the "Reset Summary Data" command in the Task menu. See Appendix 2 for details of these summaries. The default suffix used for HTML summaries is ".html" and for

text summaries is ".txt".

A LogDoor task also periodically writes out a summary report of its processing on an overall basis. This summary is written into the folder called "Overall" (or through an alias by that name) within the LogDoor output folder. The HTML overall summary report is named using the format overall<suffix>, and the text summary using the format overall<suffix>.

LogDoor's HTML summary reports can be customized by placing a text file named "Heading.html" in the same folder as the task file, and the text-based summary reports can be customized in the same way with a file name "Heading.txt". See Appendix 2 for details of these files.

Finally, you can add graphs to LogDoor's HTML summary reports to provide a quick visual interpretation of your data. Each Hits column will show the number of hits and a bar representing that number. This is implemented using a GIF-format graphics file that you supply. You tell LogDoor where this file is as follows:

- Choose "Output Files..." from the Task menu.
- Verify that "HTML Summaries" is checked.
- Enter a URL pointing to the GIF file in the "URL for Line Graph" text box. Normally this URL will be a relative URL to a GIF file on the same server as the reports, such as '/images/line.gif' but it can in fact be any valid URL.

If a GIF file is not indicated, the graphs are not drawn in the HTML summary report. A sample GIF file called line.gif is included with the software.

- IMPORTANT: All of LogDoor's summary reports are derived from the state information stored in the task file. This information reflects the data that has been processed by the task since that information was last reset, and will not necessarily reflect the information in the individual output log files (which could come from other tasks and is never reset).

LogDoor Input files

The input file for a LogDoor task can be any valid WebSTAR, AppleShare IP or other WebSTAR-compatible log file, or a HomeDoor log file. WebSTAR log files should begin with the standard format line written by WebSTAR when it starts a new log file or changes its log file format. If a LogDoor input file does not begin with a format line, the LogDoor task will derive the format of that file based on the first entry in the file.

Whenever a new input file is selected for a LogDoor task, the task will automatically configure itself to start processing at the beginning of that input file. If a task has already processed some of the current input file, but you wish to have the task restart processing at the beginning of that file, simply re-select the input file using the Input Log File button or Input File menu item.

A LogDoor task will process its input file until it has reached the end of that file. Once it has reached the file's end, the task will continue scanning that file for any new data which may be added, and it will automatically process that data. In this way LogDoor can serve as a real-time monitor of your server's activity.

- NOTE: LogDoor's input file name is stored within a task as an alias to that file. If you change the name of the file itself, or move that file, the task should still be able to locate the file. If you delete the file, however, and create another one with the same name, you will need to reselect the file.

- NOTE: With AppleShare IP, LogDoor will automatically find the input log file when a task is first created. This is possible because AppleShare IP creates a log file with a specific name in a specific location, while other servers do not.